

# ML<sup>4</sup>NGP

MACHINE LEARNING FOR NON GLOBULAR PROTEINS

## **3<sup>rd</sup> MEETING on MACHINE LEARNING AND NON- GLOBULAR PROTEINS**

**20 – 23 May 2025**

**Artis Centrum Hotel, Vilnius, Lithuania**

## 3rd MEETING ON MACHINE LEARNING AND NON-GLOBULAR PROTEINS

20 – 23 May 2025, Vilnius, LITHUANIA

### Welcome message

On behalf of the COST Action CA21160 *Non-globular proteins in the era of Machine Learning* (ML4NGP) and the *Institute of Biotechnology, Life Sciences Center, Vilnius University*, we are thrilled to welcome you to Vilnius for the 3rd ML4NGP Meeting.

We're delighted to host more than 100 scientists from over 30 countries coming together to share insights, foster collaborations, and advance the rapidly evolving field of machine learning applications in non-globular protein research. Over the coming days, the program will feature keynote lectures, invited talks, short talks, flash poster presentations, and 54 posters – all showcasing the latest advances and ideas from both within and beyond our network.

We also warmly invite you to participate in our working group meeting, where we will exchange ideas on Action's mid-term scientific evaluation and shape future directions for our community.

We hope your time in Vilnius will be both scientifically enriching and personally enjoyable, with opportunities to explore the charm, history, and vibrant culture of Vilnius.

### Organizing Committee

Darius Šulskis, Vytautas Smirnovas, Mantas Žiaunys, Andrius Sakalauskas, Kamilė Mikalauskaitė and Rūta Sniečkutė from the Institute of Biotechnology, Life Sciences Center, Vilnius University, Lithuania

Core Group of ML4NGP COST Action



## Keynote speakers

Prof. Silvio Tosatto (University of Padova, Italy)

Prof. Paul Robustelli (Dartmouth College, USA)

Prof. Ora Schueler-Furman (Hebrew University of Jerusalem, Israel)

Prof. Alan M. Moses (University of Toronto, Canada)

## Registration

The registration desk will be open on Tuesday morning, May 20th, from 8:30 to 9:45.

## Name Badges

Participants must wear their name badges at the venue and during social events, lunches, and dinners.

## Internet access

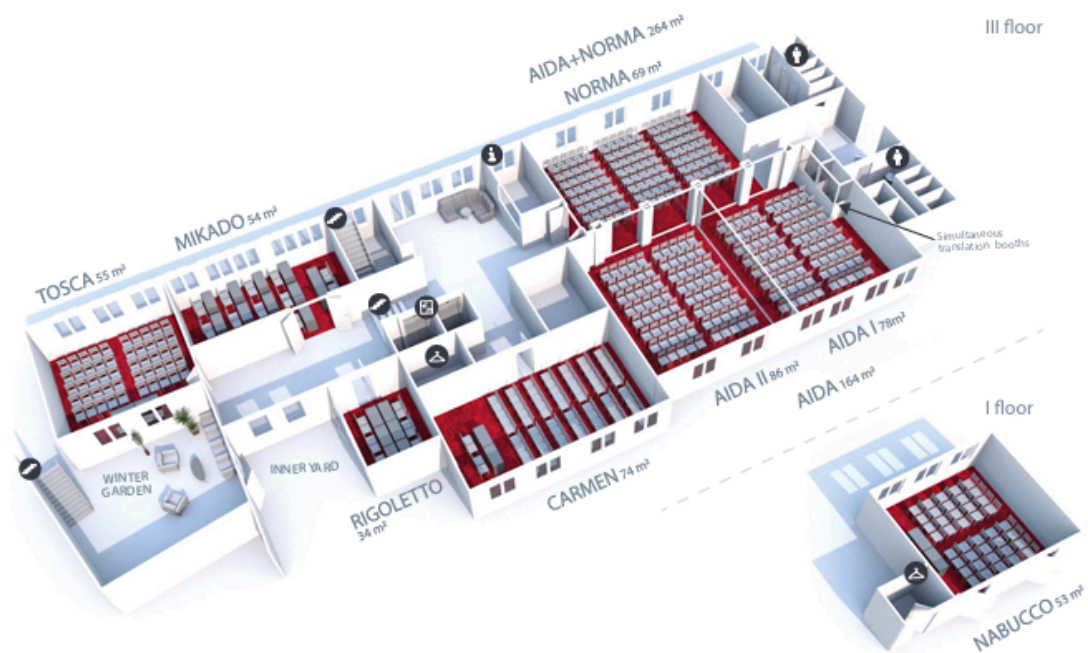
Wireless Internet is available for free.

## Venue rooms

Lectures and WG meeting will be held in the hall on the 3rd floor called AIDA. The poster presentations will be set up in the CARMEN room on Wednesday and Thursday.

Coffee breaks will be provided in the NORMA room.

Lunches will be held in the hotel restaurant on the second floor.



Location of the Conference room (Aida), poster sessions room (Carmen) and Norma (room for coffee breaks) on the Floor Plan of Artis Centrum Hotel.

## Presentation guidelines

### Poster guidelines

Posters should be up to **A0 size** – 84.1 cm (width) x 118.9 cm (height). Poster numbers will be communicated to authors at the Registration Desk and can be found in the latter chapter of this booklet.

Poster presenters are requested to set up their posters in the designated area before the poster sessions. We encourage presenters to keep their posters throughout the whole conference. Presenters are expected to be available near their posters during the assigned session. More information will be provided near the event.

### Flash Poster Presentation

Selected poster abstracts should be presented as a slide presentation (5 min). Speakers are required to have their presentation ready on a USB key and contact technical staff during the coffee break before their session.

### Oral presentation

Speakers are required to have their presentations ready on a USB key. Your slides should be either in .pdf, .ppt or .pptx format.

Speakers using their laptops are kindly asked to approach technical staff to clarify the technical prerequisites and coordinate the processes involved.

### Presentation time

Since the meeting program is fully packed, we ask all presenters to keep to the time indicated for each presentation:

Keynote speaker talks – speaking time 35 min plus 5 min discussion

Invited talks – speaking time 25 min plus 5 min discussion

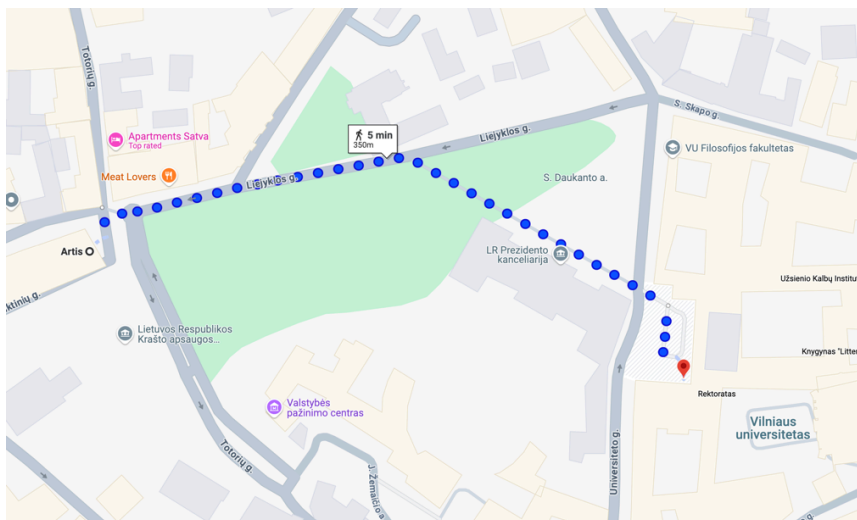
Selected talks – speaking time 12 min plus 3 min discussion

Flash presentations – speaking time 5 min

## Social event and dinner

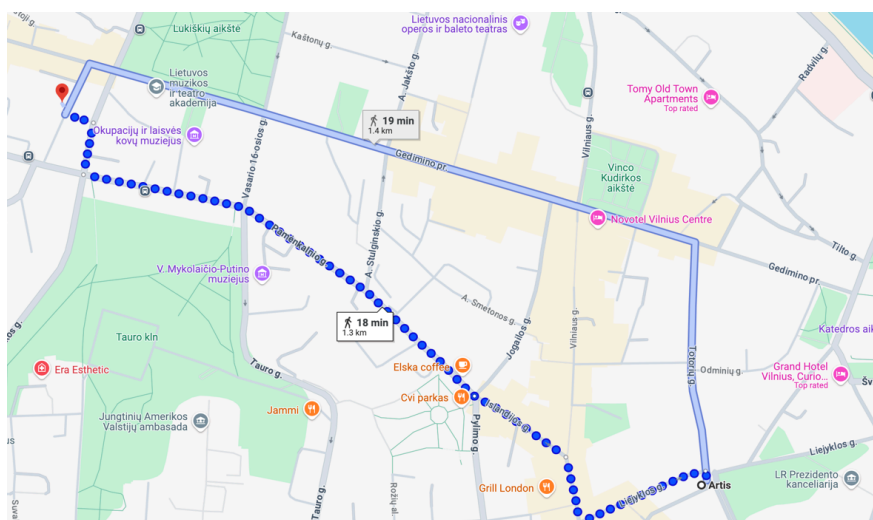
### Welcome Reception – Tuesday, 20 May 2025 | 18:00

The Welcome Reception will be held in the Aula Parva Hall (The Small Hall) at Central Vilnius University campus (<https://www.vu.lt/en/about-vu/services#aula-parva-hall-the-small-hall>). It is a 5-minute walk from the Artis Hotel. All the participants are invited to have drinks and snacks.



### Social Dinner – Thursday, 22 May 2025 | 19:00

The social dinner will be held on Thursday, May 22, 2025, at the Hilton Garden Inn (<https://www.hilton.com/en/hotels/vnogigi-hilton-garden-inn-vilnius-city-centre/>). It is a 20-minute walk from the Artis Hotel.



## Program

### **Monday, 19 May 2025**

18:00 – 19:30 Core Group Meeting (at Life Sciences Centre, Vilnius University)

### **Tuesday, 20 May 2025**

8:30 – 9:45 Registration

Conference room (Aida)

9:45 – 09:55 **Opening Session**

Alexander Monzon

» *Invited Talk*

9:55 – 10:25 **Conformational Dynamics of Intrinsically Disordered Proteins on the Surface of Gold Nanoclusters**

Pau Bernado

» *Selected Talk*

10:25 – 10:40 **On the diversity of disordered linkers in a family of multi-modular enzymes**

Juan Cortes

10:40 – 11:15 **Coffee Break**

» *Invited Talk*

11:15 – 11:45 **AI-Driven Insights into Novel Families and Functions Across the Protein Universe**

Joana Pereira

*» Selected Talks*

- 11:45 – 12:00 **A paradigmatic example of multidomain protein studied by NMR spectroscopy: The nucleoprotein of SARS-CoV-2**  
Marco Schiavina
- 12:00 – 12:15 **Maximum Entropy Method-Derived Conformational Ensembles: A FRET Perspective on a Floppy Foldase**  
Thomas-Otavio Peulen
- 12:15 – 12:30 **Sequence Features of IDRs Shape Linker Ensembles in the Context of Folded Domains**  
Hamidreza Ghafouri
- 12:30 – 14:00 **Lunch Break**  
Hotel restaurant
- 14:00 – 15:30 **ML4NGP mid-term evaluation and WGs future projects (WGs meeting)**
- 15:30 – 16:30 **Discussion and Coffee Break**

*» Selected Talks*

- 16:30 – 16:45 **Energetic Frustration as a Determinant of Intrinsic Disorder in Proteins**  
Edgar Pedro Chacon Sanchez
- 16:45 – 17:00 **Proteome-wide Prediction of the Functional Impact of Missense Variants**  
Elodie Laine
- 17:00 – 17:15 **PETIMOT: A Novel Framework for Inferring Protein Motions from Sparse Data Using SE(3)-Equivariant Graph Neural Networks**  
Sergei Grudinin
- 18:00 – 21:00 **Welcome reception**  
Aula Parva Hall at Vilnius University

**Wednesday, 21 May 2025**

Conference room (Aida)

» *Keynote lecture*

09:00 – 9:40 **Characterizing the Conformational Ensembles and Molecular Recognition Mechanisms of Intrinsically Disordered Proteins with Machine Learning**

Paul Robustelli

» *Invited Talk*

9:40 – 10:10 **Atomic resolution ensembles of Intrinsically Disordered Proteins with AlphaFold-2**

Massimiliano Bonomi

» *Selected Talk*

10:10 – 10:25 **Update on AlphaFold problems for non-globular proteins**

Dirk Linke

10:25 – 11:00 **Coffee Break**

» *Invited Talk*

11:00 – 11:30 **From accurate FRET studies on proteins to systematic assay design**

Thorben Cordes

11:30 – 12:30 **Flash poster presentations**

12:30 – 14:00 **Lunch Break**

Hotel restaurant

» *Keynote lecture*

14:00 – 14:40 **Unsupervised learning for non-globular proteins: everything I do gonna be zero-shot from now on**

Alan M. Moses

» *Invited Talks*

14:40 – 15:10 **Studying intrinsically disordered proteins by NMR and single molecule FRET spectroscopy**

Sigrid Milles

15:10 – 15:40 **Control of the dynamic self-organization of disordered proteins by post-translational phosphorylation**

Lukas Stelzl

» *Selected Talk*

15:40 – 15:55 **Molecular design features underlying nucleocytoplasmic transport**

Barak Raveh

15:55 – 16:30 **Coffee Break**

16:30 – 18:00 **Poster Session I**

**Thursday, 22 May 2025**

Conference room (Aida)

» *Keynote lecture*

09:00 – 9:40 **Confident modeling of non-confident interactions**

Ora Schueler-Furman

» *Invited Talk*

9:40 – 10:10 **New protein fibers from the dark proteome**

Andrei Lupas

» *Selected Talk*

10:10 – 10:25 **Repertoire of Side-chain Propelled Rings (SPuRs) in the C-terminal part of tau protein involved in aggregation**

Rostislav Skrabana

10:25 – 11:00 **Coffee Break**

» *Invited Talk*

11:00 – 11:30 **TBA**

Michele Vendruscolo

» *Selected Talks*

11:30 – 11:45 **Uncovering the effects of distinct membrane physical properties in tau binding to lipid vesicles and aggregation at the single-molecule level**

Ana M. Melo

11:45 – 12:00 **Can AI and ML tackle the Configurational Integral of Tau Proteins? Or actually all proteins?**

Radovan Dvorsky

12:00 – 12:15 **Structural Insights into Alpha-Synuclein Amyloid Fibril Polymorphism via Cryo-EM**

Darius Sulskis

12:15 – 12:30 **Investigating the architecture and dynamics of Cdc48-cofactor interactions through structural mass spectrometry and integrative modeling**

Tal Oppenheim

12:30 – 14:00 **Lunch Break**

Hotel restaurant

» *Invited Talks*

14:00 – 14:30 **Virulent and antimicrobial amyloids**

Meytal Landau

14:30 – 15:00 **Fuzzy interactions of IDPs driving biomolecular condensation**

Peter Tompa

15:00 – 16:00 **Round Table “Machine Learning Beyond Globule: Bridging Academia and Industry in Non-Globular Protein Research”**

Sandro Bottaro (Peptone), Peter Tompa (New Equilibrium Biosciences), Jordi Pujols (Eureka Nanobioengineering), Douglas Pires (Xyme), Silvio Tosatto (ELIXIR-IDP Community)

16:00 – 16:30 **Coffee break**

16:30 – 18:00 **Poster Session II**

19:00 – 22:00 **Social dinner**

## **Friday, 23 May 2025**

Conference room (Aida)

» *Keynote lecture*

09:30 – 10:10 **CAID-3 and beyond**

Silvio Tosatto

» *Invited Talk*

10:10 – 10:40 **SDM: Knowledge-based protein mutant stability predictor**

Arun Prasad Pandurangan

10:40 – 11:10 **Coffee Break**

» *Selected Talks*

11:10 – 11:25 **Generating temperature-dependent structural ensembles of proteins with machine learning**

Giacomo Janson

11:25 – 11:40 **Scanning coiled-coil fibers piecewise reveals local folding propensities**

Mikel Martinez Goikoetxea

11:40 – 11:55 **LCRPlatform: A Comprehensive Metaserver for Functional and Evolutionary Exploration of Low-Complexity Regions**

Aleksandra Gruca

11:55 – 13:30 **Lunch Break**

Hotel restaurant

» *Invited Talk*

13:30 – 14:00 **Dynamic ensembles of SARS-CoV-2 N-protein reveal head-to-head coiled-coil-driven oligomerization and phase separation**

Tiago Cordeiro

» *Selected Talks*

14:00 – 14:15 **GeomeTRe: A Python package for accurate estimation of geometrical parameters of Structured Tandem Repeat Proteins (STRPs)**

Zarifa Osmanli

14:15 – 14:30 **Integrative analysis of NMR and SAXS data aided with MD simulations to reveal dynamic structural features in the single alpha-helix of Drebrin**

Soma Varga

14:30 – 14:45 **Exploring the Dynamics of Non-globular Proteins: a Deep Learning Approach to Coarse-Grained Force Fields**

Andrea Guljas

14:45 – 15:00 **Closing Remarks**

**List of flash presentations:**

- F01 Oriol Bárcenas: StriFi: Comprehensive database of Fibril Structures
- F02 Gabor Erdos: AIUPred - Binding: energy embedding to identify disordered binding regions
- F03 Valentin Gonay: LLPSight: Machine learning method for prediction of LLPS driver proteins
- F04 Ailie Marx: Disordered by Design - Generative Modelling for Missing Segments in Protein Crystal Structures
- F05 Mahta Mehdiabadi: MOBIDB in 2025: integrating ensemble properties and function annotations for intrinsically disordered proteins
- F06 Samuli Ollila: Towards a Quality-Evaluated Simulation Databank for Intrinsically Disordered Proteins
- F07 Matyas Pajkos: AFlecto: A web server to generate conformational ensembles of flexible proteins from AlphaFold model
- F08 Andrius Sakalauskas: Aggregation of neurodegeneration-related proteins in artificial cerebrospinal fluid
- F09 Yuhan Wang Wang: Predictive Model Building for Aggregation Kinetics Based on Molecular Dynamics Simulations of an Antibody Fragment
- F10 Diego Javier Zea: Multiple Sequence Alignment Methods for Intrinsically Disordered Regions

**List of Poster Presentations:****Poster session I**

- F01 Oriol Bárcenas: StriFi: Comprehensive database of Fibril Structures
- F02 Gabor Erdos: AIUPred - Binding: energy embedding to identify disordered binding regions
- F03 Valentin Gonay: LLPSight: Machine learning method for prediction of LLPS driver proteins
- F04 Ailie Marx: Disordered by Design - Generative Modelling for Missing Segments in Protein Crystal Structures
- F05 Mahta Mehdiabadi: MOBIDB in 2025: integrating ensemble properties and function annotations for intrinsically disordered proteins
- F06 Samuli Ollila: Towards a Quality-Evaluated Simulation Databank for Intrinsically Disordered Proteins
- F07 Matyas Pajkos: Afflecto: A web server to generate conformational ensembles of flexible proteins from AlphaFold model
- F08 Andrius Sakalauskas: Aggregation of neurodegeneration-related proteins in artificial cerebrospinal fluid
- F09 Yuhan Wang Wang: Predictive Model Building for Aggregation Kinetics Based on Molecular Dynamics Simulations of an Antibody Fragment
- F10 Diego Javier Zea: Multiple Sequence Alignment Methods for Intrinsically Disordered Regions
- P01 Emre Aktaş: In Silico Investigation of Deleterious Missense Non-Synonymous SNPs in Lamin A (LMNA) Protein: Functional, Structural, and Molecular Dynamics Insights
- P02 Matteo Arnaudi: An Integrative Structure-Based Approach to Link Disease-Associated Mutations to Pathogenicity
- P03 Andrea Bartolomé: Mutations in human prion-like domains: pathogenic but not always amyloidogenic
- P04 Florian Blanc: Mechanism of Ag<sup>+</sup>-induced folding of a bacterial peptide from molecular dynamics simulations
- P05 Ondrej Cehlár: Structure of DC11 Fab fragment specific for the pre-aggregation conformation of IDP tau

- P06 Nevena Ćirić: Tandem repeat units rearrangement: Implications for protein structural topology
- P07 Alisa Davtyan: Prediction of amyloid propensity from the amino acid sequences using deep learning
- P08 Omar De Bei: Exploring the Role of Putative Intrinsically Disordered Regions in Staphylococcal Hemophores Involved in Iron Acquisition
- P09 Dragana Dudic: Exploring the Role of LEA6 Proteins in Drought Tolerance of Zea mays L.: A Bioinformatics and Functional Analysis
- P10 Dário Ferreira: Comparing Statistical and Machine Learning Approaches to Modeling Temporal Disorder in Non-Globular Proteins
- P11 Marc Fornt Suñé: A Novel Structure-Driven Strategy for Parkinson's Disease Immunotherapy and Diagnostics
- P12 Alba Gonzalez Artero: NEUROVAX: A Targeted Vaccine Against Toxic  $\alpha$ -Synuclein in Parkinson's Disease
- P13 Marcin Grynberg: Functions of Low Complexity Regions in protein families
- P14 Valentín Iglesias: Liquid-liquid phase separation (LLPS) as a sensing and adaptation mechanism: An evidence-based hypothesis on AP2 transcription factors in the malaria parasite
- P15 Agnė Janonienė: Protein S100A9 in neuroinflammation: exploring its direct impact on cell toxicity and intracellular accumulation
- P16 Pavel Kadeřávek: Investigation of protein sidechain dynamics in studies of intrinsically disordered proteins
- P17 Lukas Krasauskas: Tau Protein and S100A9 Co-interaction Studies
- P18 Pablo Mier: Evolutionary perspective of the CAG/CAA interplay coding for pure polyglutamine stretches
- P19 Jacopo Mlotto: Targeting undruggable intrinsically disordered regions in childhood malignant neoplasms
- P20 Stefana Njemoga: Mapping of aggregation-prone sequences within tau protein
- P21 Kliment Olechnovic: Predicting structural heterogeneity of protein contacts from singular static conformational models

P22 Fatih Ozkaynak: Application of Chaos Analysis Methods to Structural Heterogeneity and Dynamic Behavior of Non-Globular Proteins

## Poster session II

P23 Malgorzata Adamczyk: Experimental validation of non-globular sequence functions in bacterial cell division and conjugative transfer

P24 Andrea Antosova: 4,7-Disubstituted Coumarin Derivatives as Inhibitors of Amyloid  $\beta$  Peptide Fibrillization

P25 Jan Beránek: Monte Carlo simulations of miniprotein folding sampled with an autoencoder

P26 Lorenzo Bracaglia: Discriminating ordered vs. disordered modules in multidomain proteins via NMR spectroscopy

P27 Michał Burdukiewicz: Comprehensive protein datasets and benchmarking for liquid-liquid phase separation

P28 Jarosław Chilimoniuk: Aggregating amyloid resources

P29 Justas Dapkūnas: PPI3D: A Computational Resource for Searching, Analysis and Modeling Protein-Protein, Protein-Peptide and Protein-Nucleic Acid Interactions

P30 Kemal Demirtaş: Unveiling Functional Sites in Non-Globular Proteins: Directional Allosteric Networks Through Sequence-Based Gaussian Network Model and Transfer Entropy

P31 Brankica Filipic: A Long-Term Analysis of Escherichia coli Strains from Tracheal Aspirates in a Pediatric Patient—Similarity to Pandemic ST131 Strain

P32 Milana Grbic: Assessing the Impact of Graph Embedding Refinements on IDP Classification

P33 Anila Hoda: Comprehensive in silico and phylogenetic analysis of sheep  $\alpha$ s1-casein: insights into its chemical properties, structural characteristics, and functional potential.

P34 Patryk Jarnot: Challenges in adjusting scoring matrices when comparing functional motifs with non-standard compositions

P35 Filip Kapral: Mapping Protein Language: Exploring Amino Acid Functionality through Machine Learning

- P36 Katarina Martonova: STRUCTURE MOTIFS IN DISORDERED PROTEIN TAU STABILIZED BY MONOCLONAL ANTIBODIES SPECIFIC TO TAU FILAMENT CORE
- P37 Kamile Mikalauskaite: Influence of liquid-liquid phase separation on amyloid fibril structural variability
- P38 Gabriel Moya: Making Single-Molecule FRET accessible for Dynamic Biomolecular Systems using Brick-MIC
- P39 M. Victoria Nugnes: DisProt: The Manually Curated Resource for Intrinsically Disordered Proteins
- P40 Ana Pantelić: Structural characterisation of recombinant LEA protein originated from Ramonda serbica
- P41 Mykyta Peka: In silico modeling of telomerase reverse transcriptase (TERT) structure
- P42 Rafayel Petrosyan: Amyloid-forming propensity prediction and further classification of toxic vs functional amyloids
- P43 Adam Polák: Modelled conformation of tau in the PHF core
- P44 Santiago Rodriguez: Deciphering the Structural and Functional Roles of Poly-Alanine Repeats: Insights from the Combined Application of NMR and Molecular Simulations
- P45 Ahmad Shahir Sadr: Machine Learning-Driven Analysis of Protein Dynamics: A Technical Showcase with Kallikrein 3 Conformations as a Case Study
- P46 Eric Schumbera: Computational investigation of the biological and sequence context of arginine-glycine rich regions in IDRs
- P47 Rūta Sniečkutė: The Impact of  $\alpha$ -synuclein Aggregates Formed in Artificial Cerebrospinal Fluid on SH-SY5Y cells
- P48 Mariia Solovianova: Development of AmyloGraph Antibody database
- P49 Vojtěch Spiwok: Noise is all you need
- P50 Guglielmo Tedeschi: ASMSA 4 MAVISp: Analysis and Sampling by Adversarial Autoencoder for Modular Structure-Based Framework to Assess Protein Variant Effects
- P51 Mattia Utichi: Unraveling the Role of Post-Translational Modifications in Autophagy: An Integrative Study of FYCO1 and OPTN Function and Regulation
- P52 Dominykas Veiveris: Diverse effects of fluorescent labels on alpha-synuclein condensate formation during liquid-liquid phase separation



- P53 Joanna Ziemska-Legięcka: How to automatically analyse functions of Low Complexity Regions in proteins
- P54 Mantas Ziaunys: THE ROLE OF LIQUID-LIQUID PHASE SEPARATION IN ALPHA-SYNUCLEIN AMYLOID AGGREGATION

## Event Venue

The Meeting will take place at the **Artis Centrum Hotel**.

Address: Totorių street 23 LT-01120, Vilnius, Lithuania

GPS coordinates: 54.683525543133506, 25.282478540049773

Google map link: <https://maps.app.goo.gl/6uyHINjP9QvChdC27>

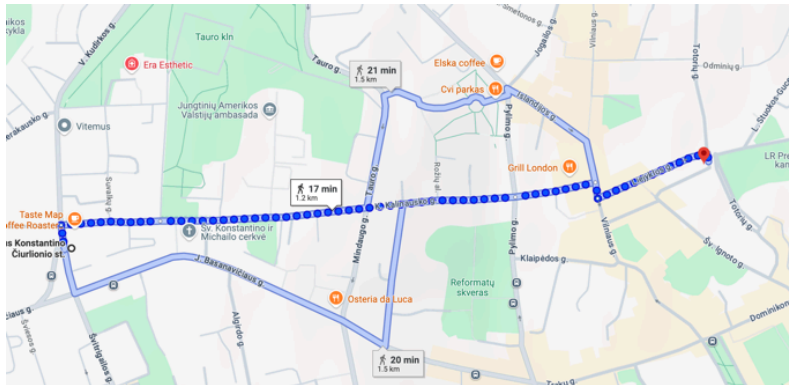
## Arriving at the hotel from the Airport

*Public transportation:*

The airport is within the city limits, take the 3G bus ( 60 min, a 0.9 EUR ticket can be bought in the bus by contactless payment card using composters or in “m.Ticket”, “Trafi” apps, more detailed explanation at <https://judu.lt/en/e-ticket-system/> website). Leave the bus at Mikalojaus Konstantino Čiurlionio st. stop, and go straight down M. Čiurlionio street until you reach the hotel.

Google map link: <https://maps.app.goo.gl/eCxsNs4ya62ZtXP9>





## Organization



*This event is part of the activities of the COST Action ML4NGP, CA21160, supported by COST (European Cooperation in Science and Technology).*

## Partner



## Sponsors

